

Pytorch - עבודת בית מס 2: ניסויים ב-XOR ובהיפר פארמטרים -

יהב יהושוע בריח (326295417) ואיתי יהודה בנדרסקי (326621927)

פלט הניסויים:

הדפס לכל אחד מהניסויים של הפרמטרים שהגדירו אותו ואת כל הערכים המחושבים בסעיף ד' המתאימים לו (מספר הפעמים שההרצה כשלה וממוצע ואחוז סטיית תקן של כל אחד מהבאים: מספר האיפוקים עד לעצירה, שגיאת אימון, שגיאת ולידציה).

```
Experiment number 0 - Learning rate: 0.01, Hidden: 2, Bypass: True
Mean epochs: 13333.8, Standard deviation epochs: 1684.405402508553
Mean train loss: 0.07729289159178734, Standard deviation train loss: 0.01135873952541818
Mean validation loss: 0.10266209244728089, Standard deviation validation loss: 0.018032208498960606
Num of fail runs: 0
```

```
Experiment number 1 - Learning rate: 0.01, Hidden: 2, Bypass: False
Mean epochs: 15264.7, Standard deviation epochs: 2452.3479382012656
Mean train loss: 0.06222695223987103, Standard deviation train loss: 0.010015669145934741
Mean validation loss: 0.07522277981042862, Standard deviation validation loss: 0.010844249175084877
Num of fail runs: 3
```

```
Experiment number 2 - Learning rate: 0.01, Hidden: 4, Bypass: True
Mean epochs: 11810.5, Standard deviation epochs: 1476.6053128713847
Mean train loss: 0.05534413270652294, Standard deviation train loss: 0.007517776328152572
Mean validation loss: 0.075787353515625, Standard deviation validation loss: 0.006669344149064327
Num of fail runs: 0
```

```
Experiment number 3 - Learning rate: 0.01, Hidden: 4, Bypass: False
Mean epochs: 12419.3, Standard deviation epochs: 1933.6359041970647
Mean train loss: 0.048408927768468855, Standard deviation train loss: 0.0035642486680616692
Mean validation loss: 0.06602277755737304, Standard deviation validation loss: 0.00331775707505608
Num of fail runs: 0
```

```
Experiment number 4 - Learning rate: 0.1, Hidden: 2, Bypass: True
Mean epochs: 2938.9, Standard deviation epochs: 240.24714358343576
Mean train loss: 0.019914943724870682, Standard deviation train loss: 0.0035645052740596755
Mean validation loss: 0.02996706683188677, Standard deviation validation loss: 0.0068283339563452354
Num of fail runs: 0
```

```
Experiment number 5 - Learning rate: 0.1, Hidden: 2, Bypass: False
Mean epochs: 2816.3, Standard deviation epochs: 712.223988644022
Mean train loss: 0.01594225037842989, Standard deviation train loss: 0.0019430318989587894
Mean validation loss: 0.021709497831761838, Standard deviation validation loss: 0.0016868573310370235
Num of fail runs: 7
```

Experiment number 6 - Learning rate: 0.1, Hidden: 4, Bypass: True
Mean epochs: 2690.7, Standard deviation epochs: 234.51953010357155
Mean train loss: 0.014713042508810758, Standard deviation train loss: 0.0015966915893163933
Mean validation loss: 0.025442792288959026, Standard deviation validation loss: 0.00364774551883185
Num of fail runs: 0

Experiment number 7 - Learning rate: 0.1, Hidden: 4, Bypass: False
Mean epochs: 2404.0, Standard deviation epochs: 351.19140080588534
Mean train loss: 0.013586723431944846, Standard deviation train loss: 0.0015916343761885427
Mean validation loss: 0.020219425670802593, Standard deviation validation loss: 0.002152805196879853
Num of fail runs: 0

Experiment number 8 - Learning rate: 0.01, Hidden: 1, Bypass: True
Mean epochs: 16634.6, Standard deviation epochs: 5179.27931666173
Mean train loss: 0.10083043947815895, Standard deviation train loss: 0.01407405561700969
Mean validation loss: 0.1254117839038372, Standard deviation validation loss: 0.0111005343271069
Num of fail runs: 1

פלט הנוירון הנסתר:

הדפס פלט הנוירון הנסתר לצד ערכי הקלט על סט האימון (עבור הניסוי מסעיף 4 בעבודה)-

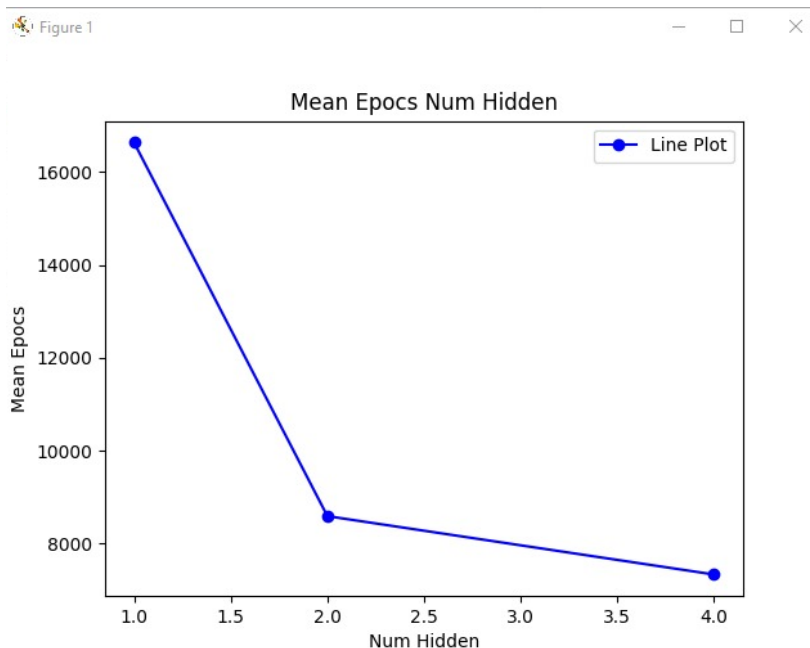
x1	x2	Output
0.0	0.0	0.061752837151288986
0.0	1.0	0.9809715151786804
1.0	0.0	0.9809406399726868
1.0	1.0	0.9999752044677734

ניתן לראות לפי טבלת האמת כי הנוירון מתפקד כפונקציה לוגית OR. זאת בכך שלפי הפלטים שהנוירון מייצר: עבור ערכי קלט כששניהם שווים ל-0 ($[0.0, 0.0]$) - הפלט קרוב ל-0 (0.0617), עבור ערכי קלט שבהם אחד המשתנים שווה ל-1 ($[1.0, 0.0]$, $[0.0, 1.0]$) - הפלט כמעט שווה ל-1 (0.9809) ועבור ערכי הקלט כששניהם שווים ל-1 ($[1.0, 1.0]$) - הפלט גם כן כמעט שווה ל-1 (0.9999). הדבר מעיד על כך שהנוירון למד להתנהג באופן הדומה לפונקציה לוגית OR, שבה מספיק שאחת הכניסות תהיה 1 כדי שהפלט יהיה 1. התנהגות זו מושגת באמצעות משקלים ו-bias המותאמים כך שפונקציית האקטיבציה (סיגמואיד) ממפה את התוצאה לטווח שבין 0 ל-1.

*ניתן לראות שבקוד הדפסתי את טבלת האמת של הנוירון הנסתר בכל הרצה בניסוי של סעיף 4, ובחרתי את אחת הטבלאות שבה ניתן לראות תפקוד דומה לפונקציה לוגית OR (אותה הטבלה שהתמונה שלה מוצגת למעלה).

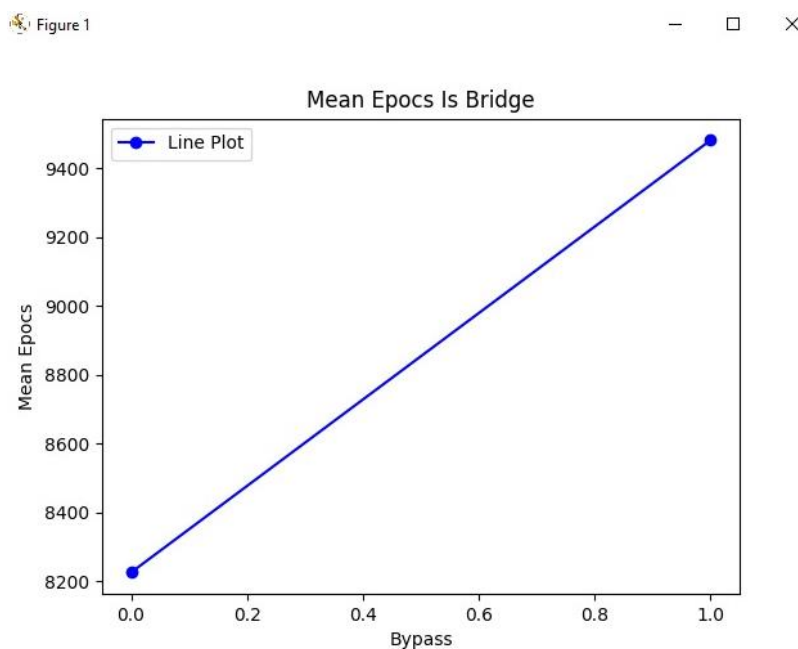
גרפים:

גרף ממוצע האיפוקים עד עצירה לעומת מספר היחידות הנסתרות –



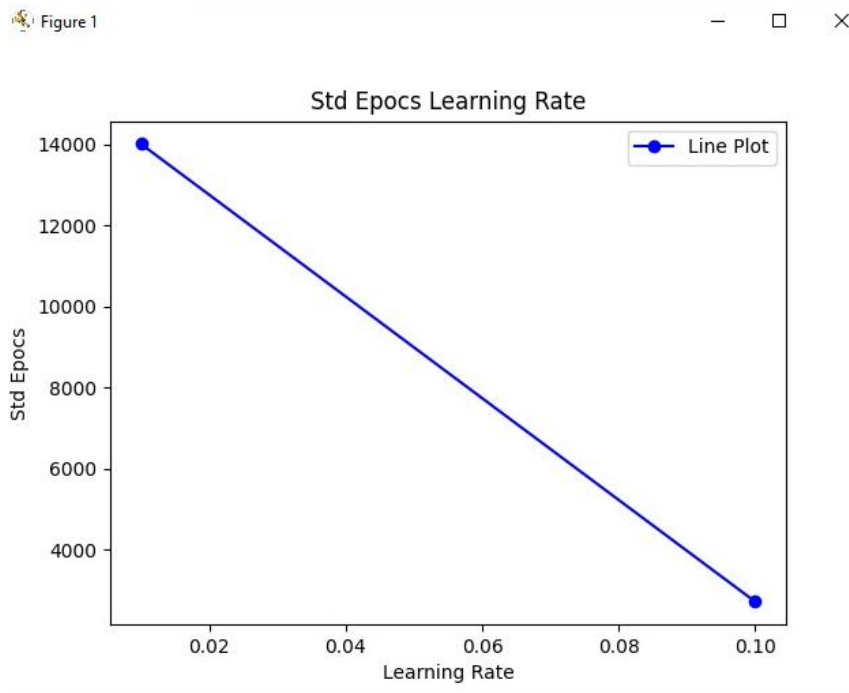
ניתן לראות כי ככל שמספר היחידות הנסתרות גדל, אז ממוצע האיפוקים קטן. זאת מכיוון שמספר יחידות נסתרות גדול יותר מגדיל את כושר הביטוי של הרשת, כלומר, הרשת מסוגלת לייצג פונקציות מורכבות יותר וללמוד ביעילות רבה יותר. כתוצאה מכך, כמות האיפוקים הנדרשת להתכנסות קטנה.

גרף ממוצע האיפוקים עד עצירה לעומת קיומו של bridge –



ניתן לראות כי כאשר bridge קיים ממוצע האיפוקים גדול יותר מאשר במצב שבו bridge אינו קיים. זאת מכיוון שהbridge יוצר חיבור ישיר בין קלט לפלט – דבר העשוי להקשות על הרשת ללמוד דפוסים מסוימים ולגרום לה לדרוש יותר איטרציות להתכנסות.

גרף סטיית התקן של מספר האיפוקים לעומת ערך learning rate –



ניתן לראות כי ככל שערך learning rate גדל, אז סטיית התקן של מספר האיפוקים קטן. זאת מכיוון ש learning rate גבוה גורם לעדכוני משקלים גדולים יותר, ולכן הרשת מתכנסת מהר יותר ובצורה עקבית יותר.